Non-Verbal Communication

Assignment 1

Group 12

Isabeau Ladage

SN: 2045940

ANR: u826443

Method & Analysis: Acoustics 1 , PRAAT

Introduction and Theoretical Background

The novel coronavirus (COVID-19) that appeared in December 2019 in Wuhan (China) has spread all over the world and has resulted in a substantial number of deaths (ECDC, 2020). Although COVID-19 involves mostly mild infections among the majority of the general population, the risk of death among young adults is higher than that of a seasonal influenza, and people with underlying comorbidities, and elderly require additional care (Kobayashi, Jung, Linton, Kinoshita, Hayashi, Miyama, ... & Suzuki, 2020).

To be able to manage the capacity of the intensive care departments of hospitals it is important that the government and citizens of countries take their responsibility. Therefore, a lot of politicians and doctors are informing their residents via national media platforms regarding procedures and techniques for high-risk people, medical staff, and the public (Saliou, 1994). Even though the risk communication is based on informing the public truthfully regarding the risk of the pandemic rather than spreading fear, the way in which the risks are communicated varies globally (Abraham, 2010). This variation is mostly non-verbal because countries all experience the virus as dangerous and warn their audiences verbally. These non-verbal differences can be explained by the cultural dimensions of Hofstede (2010), which suggest that diverse cultural values play huge roles in society (Sagiv & Schwartz, 2007; Hofstede, 2010). One important dimension which could explain possible differences between cultures is the dimension of masculinity. A high score on this dimension (Masculine) indicates that the society will be driven by competition, achievement and success, with success being defined by the best in the field. In contrast to a high score, a low score on the dimension (Feminine) means that the dominant values in society are caring for others and quality of life. A Feminine society is one where quality of life is the sign of success and standing out from the crowd is not admirable.

This suggests that in a feminine country politicians are part of society and not expected to act like they are more important or different than the majority of the population. Based on this we could expect that politicians in a feminine country are expected to show their emotions (as the residents also experience emotions), whereas politicians in a masculine country are expected to act as if they take precedence over their compatriots, pretending to be superior.

All in all, we can expect that politicians in masculine countries make less use of emotions (don't show their fears or feelings) and will therefore show less differences in voice loudness and pitch during their pitches about COVID-19 (as this suggests showing emotions), while politicians in feminine countries will make more use of emotions (show fears, anxiety or feelings) and will show more difference in voice loudness and pitch during their pitches about COVID-19, since this disease has a huge emotional impact on societies.

Based on this the following research question and hypothesis are proposed:

How do leaders in different cultures (masculine versus feminine) differ in regard to how stable their voice is in press conferences during pandemic outbreaks such as COVID-19?

Hypothesis 1 Politicians in feminine countries will use more differences in voice pitch during their speeches about COVID-19 than politicians in masculine countries

Hypothesis 2 Politicians in feminine countries will use more voice breaks during their speeches about COVID-19 than politicians in masculine countries

Hypothesis 3 Politicians in feminine countries will show a higher amount of shimmer during their speeches about COVID-19 than politicians in masculine countries

Hypothesis 4 Politicians in feminine countries will overall show more emotions by using their voice during their speeches about COVID -19 than politicians in masculine countries

Method

Data Selection

To analyze the differences in voice stableness between feminine and masculine countries 40 videos were analyzed. First, there has been a search for five countries that have the highest score on the index of masculinity by Hofstede that goes from 1 to 100. The following countries have been selected: Slovakia (100), Japan (95), Austria (79), Italy (70) and Mexico (69). Second, five countries have been selected that have the lowest score on the index of masculinity and are therefore more feminine type of countries: Sweden (5), Norway (8), Latvia (9), Netherlands (14) and Denmark (16) (Hofstede, 2011). For every country four videos have been selected via the platform YouTube. The video material all consisted of the minister or president of the country in a press conference regarding the coronavirus or videos in which an official statement on television regarding the coronavirus was made. Moreover, videos were selected where ministers or presidents talked in their own native language. All selected videos were uploaded between the 4th of March and the 24th of April, and the links of these video's can be found in Appendix 1. Therefore the videos go back to different dates and represent press conferences in different times during the corona crisis. The videos were 1080p50 HD. The recording situation was the same in each video, namely front angle. There were no background noises during speaking, the ministers or presidents articulated good and did not mumble.

Acoustics 1

Materials

First, the 40 different videos found on YouTube were converted into files (wav.), since only the audio is needed. Next, for every file the first minute from the moment the minister starts talking is selected.

This results in 40 minutes of audio material which is named by stating the country in the title. Hereby the audio fragments do not contain voices of any other people such as journalists and for example no intro tunes. Analyzing the whole video would have made the dataset less easy to compare.

Measurements

The goal of this method is to measure the acoustics. To do this there is made use of an open-source program for the analysis of speech in phonetics: 'PRAAT'. All 40 sound files were uploaded to this tool which is more reliable than recording the videos straight into PRAAT. To measure voice stableness, for every sound file the following measurements are extracted: Pitch (Max-min), voice breaks, and shimmer. The dataset has then been uploaded in SPSS to do the statistical analysis.

Through the option Voice Report, minimum and maximum Pitch height in Herz was being measured by the programm. Manually the pitch floor is subtracted by the pitch ceiling, which results in a score for the variation in pitch. The Voice Report in PRAAT also provided a total number of voice breaks in the fragment and a percentage of Shimmer (local). The amount of voice breaks is the number of distances between consecutive pulses that are longer than 1.25 divided by the pitch floor (Boersma & Weenink, 2003). Shimmer measures the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude (Boersma & Weenink, 2003).

Finally, for every country the average of these data is calculated and the countries are divided into two groups: Feminine countries and Masculine Countries (F & M). This provided us with the following averages of the three acoustic measurements (See table 1).

F	149,12	134,30	12,19
M	161,08	133,65	12,39

Table 1. *Means of Pitch variation, Voice breaks and shimmer*.

Perception

Participants and design

In total, 56 participants took part in the study, of which 33 were female (58.9%) and 23 were male (41.1%). Of the participants, one had the Belgium nationality, all others had a Dutch nationality. The youngest participant was 18 years old and the oldest was 56 years old (M = 25.04, SD = 7.90). 91% of the participants were higher educated (HBO, University Bachelor, University Master), and 9% were lower educated (High School or MBO).

The effect of feminine countries compared to masculine countries (independent variable) on the voice stableness of their presidents during press conferences about the corona crisis situation (dependent variable) was studied. This was done by perception study with a between-subjects online experiment, in

which participants were randomly assigned to one of two conditions in which participants both got to hear 20 sound fragments from ten different countries (2 video's per country). So, in total 40 sound fragments were used. On the one hand, 20 sound fragments were related to masculine countries. On the other hand, 20 fragments were related to feminine countries. The links of the video clips can be found in Appendix 1. However, participants only got to hear 20 sound fragments, 10 masculine fragments and 10 feminine fragments each. Each fragment only lasted for fifteen seconds to avoid participants becoming fatigued or mind wandering. Shimmer, pitch variation, and voice breaks were measured to be able to determine the voice stableness in the end.

Materials

The stimuli in the experiment consisted of parts of Youtube videos. All these videos consisted of speeches related to the current COVID-19 crisis and are held by preministers. For equal comparison, these videos were cut and only the first 15 seconds after the president or leader of the country started to speak were put into .m4a sound files (as after 15 seconds some leaders were interrupted). By doing so it was doable for participants to listen to 20 fragments and keep their concentration and focus. The video's could be paused and rewound by the participant. Participants saw immediately after each sound fragments the questions they had to answer about the corresponding fragment. These questions were translated into Dutch, to avoid translation bias from the participants themselves. The questions and their translation can be found in Appendix 2. The continue button appeared after 10 seconds to force participants to listen to the sound fragment and answer the questions related towards the fragments (10 seconds was the least estimated time participants needed to listen to the sound files). These timers were used to prevent participants from rushing through the experiment (i.e., skipping the sound files).

Measurement

Shimmer

Shimmer can be defined as the variation in the loudness of the intensity or energy of the voice (Brockmann, Drinnan, Storck & Carding, 2011; Matsumoto, Frank & Hwang, 2012). According to Verdolini and Ramig (2001), the level of shimmer can be determined by the change in loudness of the voice. More perceived shimmer indicated a more unstable voice (Brockmann et al., 2012). For measuring shimmer two questions on a 7-points likert scale, ranging from strongly disagree to strongly agree, were presented to the participants after every sound file. The two items were: "The loudness of the voice varies" and "The volume of the voice remains constant". The reliability of the shimmer scale was good ($\alpha = .87$).

Pitch variation

The term pitch is used to describe the measurement of the vocal frequency (Lieberman, 1961). Moreover, pitch refers to the perceived height of the tone of the voice (Smith, Lemke, Taylor, Kirchner

& Hoffman, 1998). More pitch variation indicated a more unstable voice (Smith et al., 1998). For measuring the pitch variation two questions on a 7-points likert scale, ranging from strongly disagree to strongly agree, were presented to the participants after every sound file. The two items were: “The height of the voice varies” and “The tone of the voice remains constant”. The reliability of the pitch scale was good (α = .82).

Voice breaks

Voice breaks can be defined as an intrinsic instability of the movement of the vocal folds (Berry, Herzel, Titze & Story, 1996). Voice breaks may occur during speaking when the voice suddenly switches into the falsetto register, or the other way around (Bloothooft, Wijck & Pabon, 2001). According to Bloothooft et al. (2001) these voice breaks may either occur spontaneously or voluntary, in both male and female adults. However, later in life these voice breaks rarely occur spontaneously. Voice breaks are most likely to occur in the initiation or termination of speaking (Duffy, 1970). This is another reason why the first fifteen seconds of each sound fragment were used. Voice breaks are easily perceived even by untrained listeners (Colton, Casper & Leonard, 2011). For measuring the occurrence of voice breaks two questions on a 7-points likert scale, ranging from strongly disagree to strongly agree, were presented to the participants after every sound file. The two items were: “The speaker’s voice seems to break” and “The speaker’s voice does not break”. The reliability of the comprehension scale was good (α = .80).

Control variable

To be able to ensure whether the answers of the participants on the three measurements above indeed indicated that the participants found the voice in the fragment stable or not a control item was added. This control item consisted of two questions on a 7-points likert scale, ranging from strongly disagree to strongly agree, which were presented to the participants after every sound file. Afterwards, the average scores of the three items and the control item were compared to find similarities. Similarities between these scores indicated that the participant’s answers on the three measurements indeed indicated that the participants found the voice stable or not. The two items were: “I think this voice is stable” and “I do not think this is a stable voice”. The reliability of the comprehension scale was good (α = .81).

Procedure

Convenience sampling was used to recruit participants. Participants were sent a digital invitation to the online experiment. Before commencing with the experiment, informed consent was asked. Participants were instructed about the experiment, asked for their time, and given the option to start by continuing.

The experiment started with some information for the participants, which were told that they were going to be exposed to several sound fragments in different languages. The participants were told that it didn’t matter if they didn’t understand what was said, and were asked to listen to the fragments

and answer the scale questions below the fragment. Finally, demographic questions were asked and the participants were debriefed.

Acoustics 2

Materials

Firstly, the 40 different videos were selected from YouTube to analyze vocal stability differences between feminine and masculine countries. 20 videos consist of speeches from most masculine countries from the masculinity index. The remaining 20 videos consist of speeches from feminine countries which have low scores in masculinity index.

Since the main purpose of this method is to measure the acoustics in the context of vocal stableness, all selected video samples were downloaded from YouTube and converted into wav. files For every video, the first minute the prime minister in the video actively starts talking about coronavirus is selected and collected for further analysis. Every video was recorded and uploaded to YouTube after February 2020. After the conversion, all 40 sound files were uploaded into PRAAT, an open-source program regarding their types (Feminine vs. Masculine) to analyze voice stability. On the other hand, none of the audio fragments contain any other background noises or voices of journalists. Whole audio files only contain the prime minister's voice. For every sound file, the following measurements are extracted: pitch(Mean), jitter, intensity, and pauses. The dataset has then been uploaded in SPSS to do the statistical analysis.

Measurements

Pitch

Firstly, pitch is measured by Pitch mean. Praat measures and automatically provides Pitch mean based on Pitch range and Pitch height in Herz through the Voice menu option. High pitch mean is highly related to emotional intensity of the speaker (Dietrich, Hayes & O'Brian, 2019). Additionally, related to the length of vocal folds, women and children tend to have higher pitched voices than men (Daniloff, Schuckers & Feth, 1980).

Jitter

Jitter can be defined as the parameter for frequency parameters in the speech (Teixeira, Oliveira & Lopes, 2013) and identifies the variations in the pitch. Large variations generally refer to instability in the voice (Kamiloglu, Fischer & Sauter, 2020). Additionally, jitter is measured as the average absolute difference between consecutive periods, divided by the average period.

Pauses

Finally, pauses are identified as the silence in a speech between two words (Frank,Griffin, Maroulis & Svetieva, 2014). Generally, pauses are caused by the interpolarazation of the speech based on whether the speech is understandable by the listeners (Boomer,1970). Therefore, during speeches pauses refer to semantic and lexical decisions of the speaker (Boomer,1970). Boomer and Dittman (1962) claim that

pauses in the speech should be between two and five seconds. Therefore, the silences more than two seconds were identified as pauses and have been counted manually through the View options of Praat.

Intensity

Intensity, on the other hand, refers to the measurement of loudness (Frank,Griffin, Maroulis & Svetieva, 2014). It is measured by the mean of the intensity values of the frames through Praat. It is highly related to the emotional status of the speaker. Intensity measurements can identify the differences between low and high arousal emotions. (Goudbeek & Scherer, 2010).

Finally, the mean of these data is calculated and are divided into country type: Feminine countries and Masculine countries. This provided us with the following statistical analysis of the four acoustic measurements

Results

Acoustics 1

First, the correlation coefficients for the three variables are calculated with Pearson's r. There is no correlation found between Pitch Variation and Voice Breaks (r=.124, p>.445). For Pitch Variation and Shimmer the Pearson-correlation did not show a significant correlation either (r=.092, p>.861). However, Shimmer does seem to significantly correlate with Voice Breaks (r=.496, p<.001)**.** The strength of this correlation can be defined as moderate and has a positive direction. The correlation between Shimmer and Voice Breaks is made more explicit with a scatter plot in Figure 1.

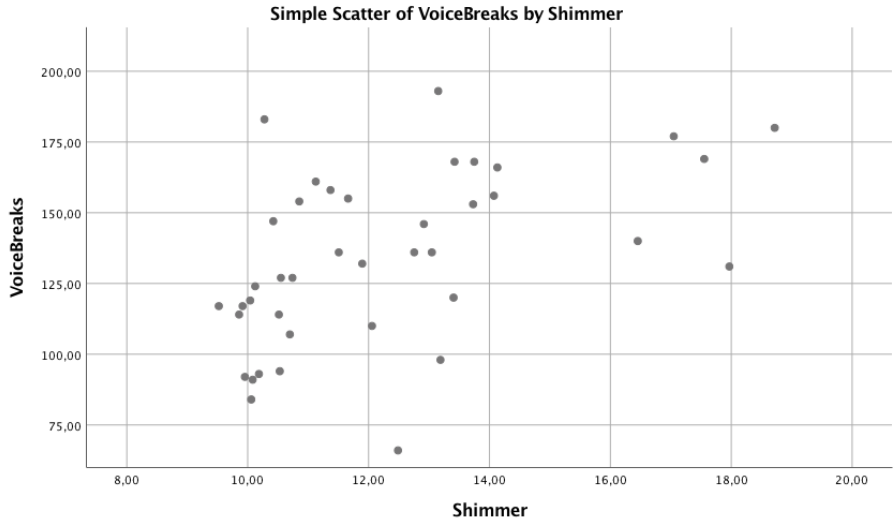

Figure 1. *Correlation Voice breaks and Shimmer*.

Next, three independent t-tests were conducted to compare the three variables between the two conditions: feminine and masculine countries.

From the t-tests can be concluded that there is a significant difference in Pitch Variation between the feminine (*M*=149,12; *SD*=12,34) and masculine (*M*=161,08; *SD*=7,08) countries (*t*(38)=-3,76, *p*<.001). This means that speakers of masculine countries show more variation in pitch height during the press conferences and therefore talk with a less stable voice. This contradicts Hypothesis 1, where it is expected that the group of feminine countries would show more variation in Pitch.

The average amount of voice breaks found in the videos of feminine countries (*M*=134,30; *SD*=26,09) is slightly higher than masculine countries (*M*=133,65; *SD*=35,13) but no significant difference is found here (*t*(38)=,06, *p*>.947).

For the percentage of shimmer there has again not been found a significant difference between the feminine countries (*M*=12,19; *SD*= 2,61) and masculine countries (*M*=12,39; *SD*=2,33) (*t*(38)=-,25, *p*>.802). Concluding, all three hypotheses are not being supported by these findings.

Perception

The directions of the correlations between outcome variables are expected to are highly correlated as they all measure the overall concept voice stableness. Especially, shimmer and pitch variation correlated highly, as these measurements are both related to vocal effort (Weis, 2019), which was in line with the expectations. However, the amount of voice breaks didn't correlate high with either pitch variation or the amount of shimmer. This could be explained by the fact that voice breaks are easily perceived by untrained listeners (Colton, Casper & Leonard, 2011), whereas pitch variation and shimmer may be less easily perceived by untrained listeners. Lastly, as expected shimmer, pitch variation and voice breaks correlate negatively with voice stableness. This is probably due to the fact that a high score on one of these three aspects refers to a low score on stableness. The corresponding correlation scores are presented in Table 2 below.

Table 2 *Full correlation matrix*

Correlations					
		Shimmer	Pitch	VoiceBreak	Stableness
Shimmer	Pearson Correlation	1	,494**	,256	-,268*
	Sig. (2-tailed)		,000	,057	,045
	N	56	56	56	56
Pitch	Pearson Correlation	,494**	1	,099	-,202
	Sig. (2-tailed)	,000		,469	,136
	N	56	56	56	56
VoiceBreak	Pearson Correlation	,256	,099	1	-,550**
	Sig. (2-tailed)	,057	,469		,000
	N	56	56	56	56
Stableness	Pearson Correlation	-,268*	-,202	-,550**	1
	Sig. (2-tailed)	,045	,136	,000	
	N	56	56	56	56
* / **. Correlation is significant at the 0.05 / 0.01 level (2-tailed).					

To investigate whether one type of culture (masculine versus feminine) results in more shimmer, more pitch variation, more voice breaks, and a higher voice stableness a MANOVA was performed. In this analysis “type of culture” was seen as an independent variable, whereas “shimmer”, “pitch variation”, and “voice breaks” were seen as dependent variables. “Overall voice stableness” was seen as a control variable. For “type of culture” two levels were measured: masculine and feminine. The dependent variables “shimmer”, “pitch variation” and “amount of voice breaks”, and the control variable “overall voice stableness”, were all measured on four different continuous scales ranging from 1-7.

After performing the MANOVA, the Wilks's Λ showed that there was no significant difference between the different groups $\Lambda = 0.698$, $F(4, 51) = .553$, $p = .698$. First of all, the amount of pitch variation across masculine and feminine countries was distributed normally. On average, the perceived amount of pitch variation was higher when people heard a feminine fragment ($M = 4.02$, $SD = .53$) than when they heard a masculine fragment ($M = 3.99$, $SD = .56$). Equal variance between groups was assumed $F(1, 54) = .13$, $p = .717$. The difference between pitch variation in masculine and feminine fragments was not significant $F(1, 54) = .036$. $p = .850$, eta-squared= .001. Therefore, it cannot be concluded that politicians in feminine countries show variation in pitch during their speeches about COVID-19 than politicians in masculine countries, which is why H1 is rejected.

Furthermore, the amount of voice breaks across masculine and feminine countries was distributed normally. On average, the perceived amount of shimmer was higher when people heard a feminine fragment ($M = 3.33$, $SD = .58$) than when they heard a masculine fragment ($M = 3.26$, $SD = .62$). Equal variance between groups was assumed $F(54) = .25$, $p = .619$. The difference between shimmer in masculine and feminine fragments was not significant $F(1, 54) = .189$. $p = .665$, eta-squared= .003. Thus, the H2 is rejected, it cannot be concluded that politicians in feminine countries show more voice breaks during their speeches about COVID-19 than politicians in masculine countries.

In addition, the amount of shimmer across masculine and feminine countries was distributed normally. On average, the perceived amount of shimmer was higher when people heard a masculine fragment ($M = 4.04$, $SD = .31$) than when they heard a feminine fragment ($M = 3.98$, $SD = .38$). Equal variance between groups was assumed $F(1, 54) = 1.07$, $p = .307$. Univariate tests showed that the groups (masculine versus feminine) did not differ on shimmer, as masculine $F(1, 54) = .398$. $p = .531$, eta-squared= .007. All in all, it cannot be concluded that politicians in feminine countries show more shimmer during their speeches about COVID-19 than politicians in masculine countries, which is why H3 is rejected.

Finally, the mean scores of this control variable voice stableness do not differ much between masculine ($M = 4.52$, $SD = .33$) and feminine countries ($M = 4.60$, $SD = .46$). The scores of voice stableness across masculine and feminine countries were distributed normally and equal variance between groups was assumed $F(54) = 2.12$, $p = .151$. The difference between voice stableness in masculine and feminine fragments was not significant $F(1, 54) = .453$. $p = .504$, eta-squared= .008. Accordingly, it cannot be concluded that politicians of feminine countries have a lower voice stableness

than politicians of masculine countries. All in all, none of the three aspects of voice stableness showed a significant difference between politicians of feminine and masculine counties. This is also reflected in the control variable, which is insignificant as well. Therefore, it cannot be concluded that politicians in feminine countries show more emotions in their voice during their speeches about COVID-19 than politicians in masculine countries. Hence, hypothesis 4 is rejected.

Acoustics 2

The correlation coefficients for the four variables are calculated with Pearson's *r*. A significantly correlation was found between pitch (mean) and intensity (r=-.322, p<.05). This results reveal that there is moderate and negative correlation between two variables. Because both pitch (mean) and intensity are highly related to emotional intensity, we were expecting a positive correlation related to emotional vocal expressions. On the other hand, pauses are also significantly correlated with jitter (r=-.411, p<.01). This correlation is also negative. However, this negative correlation can be explained that if there is more pitch variation cycle by cycle during the speech, the counts of the silence between words would decrease. The correlation matrix can be seen in Table 3.

		Intensity	Jitter	Pitch (Mean)	Pauses
Intensity	Pearson Correlation	1	.008	.167	-.322*
	Sig. (2-Tailed)	-	.963	.303	.043
	N	40	40	40	40
Jitter	Pearson Correlation	.008	1	-.441**	-.125
	Sig. (2-Tailed)	.963	-	.004	.443
	N	40	40	40	40

Pitch (Mean)	Pearson Correlation	.167	-.441**	1	-.60
	Sig. (2-Tailed)	.303	.004	-	.715
	N	40	40	40	40
Pauses	Pearson Correlation	-.322*	-.125	-.060	1
	Sig. (2-Tailed)	.043	.443	.715	-
	N	40	40	40	40

*. Correlation is significant at the 0.05 level (2-tailed)

**. Correlation is significant at the 0.01 level (2-tailed)

Table 3 *Correlation Matrix*

Four independent t-tests were conducted to compare the four variables between the two conditions: feminine and masculine countries.

Firstly, to test pitch(Mean) an independent T-test was performed. On average, feminine countries' pitch(Mean) results (*M*= 174.03, *SD* = 41.89) were higher than masculine countries' pitch results (*M* = 164.33, *SD* = 22.15). However, this difference was not significant (*Mdif* = 9.70, *t*(38) =0.916, *p* =0.37). The average amount of jitter was measured in the videos of feminine countries (*M*=0.021; *SD*=.008) is slightly higher than masculine countries (*M*=0.020; *SD*=.006) but no significant difference is found here (*Mdif* = .001, *t*(38)=.05, *p*=0.6). On the other hand, although the intensity results of masculine countries (*M*=68; *SD*=5,33) were little higher than the feminine countries (*M*=66.03; *SD*= 5,91), there has again not been found a significant difference (*Mdif* = -2.42, *t*(38)=-1.35, *p*=0.18). Finally, the average count of pauses were measured in the videos of feminine countries (*M*=0.06; *SD*=.006) is slightly higher than masculine countries (*M*=0.05; *SD*=.006) but no significant difference is found here (*Mdif* = .05, *t*(38)=.23, *p*=0.8).

In sum, because none of statistical results shows any significant difference, all four hypotheses are not being supported.

Conclusion & Discussion

The aim of this study was to investigate if there are differences in emotion in voice between leaders of feminine and masculine countries, during their official speeches regarding COVID-19.
The first hypothesis, that leaders of feminine countries will show more variation in their voice pitch, has not been confirmed by any of the methods.
Contradicting, the acoustic measurement even found a significant opposite effect: the leaders of masculine countries showed higher levels of Pitch variation.

Hypothesis two expected leaders of feminine countries to have more voice breaks during their speeches than leaders of masculine countries. Again, this hypothesis is not supported by our findings.

Hypothesis three can not be confirmed either, since there are no significant differences found between feminine and masculine countries regarding Shimmer.

Consequently, hypothesis four, that expects leaders of feminine countries to show more emotions by using their voice than leaders of masculine countries, can not be confirmed.

Since there are no results found, the effect of the score on masculinity of a country might be questioned. There might have been all sorts of other variables that created differences within the two categories of countries. For example the gender of the leader of the country. Both groups consisted of both female and male speakers. In previous research a significant interaction between leader gender and emotion was found (Lewis, 2000). Lewis found that female leaders in general show more emotion in voice.

Second, the emotion in voice might also had to do with the current severity of the coronavirus that the country was experiencing and the type of message the person was about to bring. Needing to bring an unpleasant message might come along with more stress and emotions than having a positive message for the citizens. Research found that voice stableness is also influenced by the level of stress a person experiences (Tao et al., 2007). This level of stress can be related to the current situation the country was in, or the news he or she had to bring. Moreover, Scherer (1986) found that 'sad' vocalizations tend to have less pitch variability and are slower. Whether the message the leader was bringing was in general more negative and sad than positive has not been taken into account.

Third, language might also be a strong indicator for voice stableness. Previous research found that intonation and rhythm can vary strongly across language families (Scherer, Banse, and Wallbott, 2001).

Lastly, the sample of data might have been too small. To really get a proper perspective on the effect of the level of masculinity of a country, videos of more countries should have been analyzed.

To conclude with, future research in this field should take into account more variables that might have an effect on voice stableness in times of pandemics. For example: gender, language and the type of message the person is bringing during the press conference. Additionally, a larger data

selection and therefore the analysis of more countries, would also give more opportunities to draw generalized conclusions. Especially the findings of Lewis (2000) give food for thought for future, more practical, research: An angry voice may create motivation to work harder to improve the situation, while expressing sadness may result in passive acceptance rather than to put in effort to make the situation less worse. This would be extremely relevant in terms of the effects of the press conferences about the corona virus on the behavior of the audience.

In order to be able to answer the research question, two methods have been used: perception measurements and acoustics. These two methods come with advantages and disadvantages.

First, previous research found that general attributes of voice quality may be distracting for foreign listeners and may overshadow the actual vocal cues used to express emotions (Elfenbein & Ambady, 2002). This could have had an impact on the perception measurements, since all participants were Dutch and therefore listening to both their native language and foreign languages.

Second, for the acoustic measurements there is made use of the programm PRAAT. In previous research there is found that different results in voice perturbation are found between different programs (Maryn et al., 2009). For example, the Multi-Dimensional Voice Program is proposed as a better alternative here. However, PRAAT does provide the researcher with a large variety of measurements. Therefore, we were able to investigate seven different measurements.

Also, the same videos are used for both methods but the length of the fragments was different. For the acoustic measurement the analyzed voice fragments had a length of one minute, while for the perception analysis the fragments were only 15 seconds long. These 15 seconds might have been a part of the speech where the speaker said something specific that made his or her emotion in voice arise.

Additionally, it would have been interested for this research to consist of a third method: expert coding. However, it was too difficult to find experts in voice for every different language. More effort should be made in future research to find native speakers in all languages to enlarge the reliability of measuring both perception and expert coding.

References

Elfenbein, H. A. and Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin* 128, 208-235.

Lewis, K. M. (2000). When leaders display emotion: How followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *21*(2), 221-234.

Li, X., Tao, J., Johnson, M. T., Soltis, J., Savage, A., Leong, K. M., & Newman, J. D. (2007, April). Stress and emotion classification using jitter and shimmer features. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07* (Vol. 4, pp. IV-1081).

Scherer. K. R. (1986). Vocal affect expression: A review and model for future research. *Psychological Bulletin* 99, 143-165.

Scherer, K. R., Banse, R., & Wallbott, H. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32,

Abraham, T. (2010). The price of poor pandemic communication. *Bmj*, *340*(jun09 2), c2952–c2952 doi: 10.1136/bmj.c2952

Berry, D. A., Herzel, H., Titze, I. R., & Story, B. H. (1996). Bifurcations in excised larynx experiments. *Journal of Voice*, *10*(2), 129–138. https://doi.org/10.1016/s0892-1997(96)80039-7

Bloothooft, G., Wijck, M.V., & Pabon, P. (2001). Relations between vocal registers in voice breaks. *INTERSPEECH*. https://www.isca-speech.org/archive/archive_papers/eurospeech_2001/e01_0039.pdf

Boersma, P. & Weenink, D. (2003, May 21). Voice 3. Shimmer. Retrieved on 16 May 2020, from http://www.fon.hum.uva.nl/praat/manual/Voice_3__Shimmer.html

Boersma, P. & Weenink, D. (2003, September 16). Voice 1. Voice breaks. Retrieved on 16 May 2020, from http://www.fon.hum.uva.nl/praat/manual/Voice_1__Voice_breaks.html

Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice*, *25*(1), 44–53. https://doi.org/10.1016/j.jvoice.2009.07.002

Boomer, D. S. (1970). Psycholinguistics; experiments in spontaneous speech. *Lingua*, *25*, 152–164. doi: 10.1016/0024-3841(70)90028-8

Colton, R. H., Casper, J. K., & Leonard, R. J. (2011). *Understanding voice problem: A physiological perspective for diagnosis and treatment: Fourth edition*. Wolters Kluwer Health Adis (ESP).

Dietrich, B. J., Hayes, M., & O'Brien, D. Z. (2019). Pitch Perfect: Vocal Pitch and the Emotional Intensity of Congressional Speech. *American Political Science Review*, *113*(4), 941–962. doi: 10.1017/s0003055419000467

Duffy, R. J. (1970). Description and Perception of Frequency Breaks (Voice Breaks) in Adolescent Female Speakers. *Language and Speech*, *13*(3), 151–161. https://doi.org/10.1177/002383097001300302

European Centre for Disease Prevention and Control. "Situation Update—Worldwide". 2020. Available online: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases (accessed on 13 February 2020).

Frank, M. G., Griffin, D. J., Svetieva, E., & Maroulis, A. (n.d.). Nonverbal Elements of the Voice. *The Social Psychology of Nonverbal Communication*. doi: 10.1057/9781137345868.0010

Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture*, *2*(1), 8.

Kamiloglu, R. G., Fischer, A., & Sauter, D. (2019). Good Vibrations: A Review of Vocal Expressions of Positive Emotions. doi: 10.31234/osf.io/86rmu

Kobayashi, T., Jung, S. M., Linton, N. M., Kinoshita, R., Hayashi, K., Miyama, T., ... & Suzuki, A. (2020). Communicating the risk of death from novel coronavirus disease (COVID-19).

Lieberman, P. (1961). Perturbations in Vocal Pitch. *The Journal of the Acoustical Society of America*, *33*(5), 597–603. https://doi.org/10.1121/1.1908736

Matsumoto, D., Frank, M. G., & Hwang, H. (2012). *Nonverbal Communication: Science and applications* (1st ed.). https://doi.org/10.4135/9781452244037

Teixeira, J. P., Oliveira, C., & Lopes, C. (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, *9*, 1112–1122. doi: 10.1016/j.protcy.2013.12.124

Sagiv, L., & Schwartz, S. H. (2007). Cultural values in organisations: insights for Europe. *European J. of International Management*, *1*(3), 176–190. https://doi.org/10.1504/ejim.2007.014692

Saliou, P. (1994). Crisis communication in the event of a flu pandemic. *European Journal of Epidemiology*, *10*(4), 515–517. doi: 10.1007/bf01719693

Smith, E., Lemke, J., Taylor, M., Kirchner, H. L., & Hoffman, H. (1998). Frequency of voice problems among teachers and other occupations. *Journal of Voice*, *12*(4), 480–488. https://doi.org/10.1016/s0892-1997(98)80057-x

Verdolini, K., & Ramig, L. O. (2001). Review: Occupational risks for voice problems. *Logopedics Phoniatrics Vocology*, *26*(1), 37–46. https://doi.org/10.1080/140154301300109125

Weiss, B. (2019). *Talker Quality in Human and Machine Interaction: Modeling the Listener's Perspective in Passive and Interactive Scenarios (T-Labs Series in Telecommunication Services)* (1st ed.). Springer Nature, Switzerland AG: Springer International Publishing. https://doi.org/10.1007/978-3-030-22769-2

Appendices

Appendix 1. *Links of video clips per country*

Masculine countries

Slovakia

1. https://www.youtube.com/watch?v=Q17ZxKXlr20
2. https://www.youtube.com/watch?v=HWBNvdUBdbQ
3. https://www.youtube.com/watch?v=rkcYQidaUnw
4. https://www.youtube.com/watch?v=BB4O6MbIdww

Japan

1. https://www.youtube.com/watch?v=tCFhv6qoAs8
2. https://www.youtube.com/watch?v=ion9Bfl6eJw
3. https://www.youtube.com/watch?v=IBW5ix1AYqA
4. https://www.youtube.com/watch?v=uV6RXV5wbi4

Austria

1. https://www.youtube.com/watch?v=UbI7Ihn2e7o
2. https://www.youtube.com/watch?v=1AwLGjevQFw
3. https://www.youtube.com/watch?v=vpxgsUwhmNY
4. https://www.youtube.com/watch?v=hKlBO7cUND0

Italy

1. https://www.youtube.com/watch?v=HxOEK76QyV0
2. https://www.youtube.com/watch?v=-J__TiMVwlg
3. https://www.youtube.com/watch?v=FO8Xg-BRqoU
4. https://www.youtube.com/watch?v=f9h-1ugJKJQ

Mexico

1. https://www.youtube.com/watch?v=Ykuk6Ur_kvY
2. https://www.youtube.com/watch?v=f9jJsPCYaYk
3. https://www.youtube.com/watch?v=Zr1kUs9F-7Y
4. https://www.youtube.com/watch?v=Lx1sk5yqwq8

Feminine countries

Sweden

1. https://www.youtube.com/watch?v=QxmsRfQxXCc
2. https://www.youtube.com/watch?v=AIX25_1tk4A
3. https://www.youtube.com/watch?v=A09tAefUfww
4. https://www.youtube.com/watch?v=GsdraUeoB1Q

Norway

1. https://www.youtube.com/watch?v=MJ1W9Un7dMo
2. https://www.youtube.com/watch?v=F48zJ9IwEqY
3. https://www.youtube.com/watch?v=MJ1W9Un7dMo
4. https://www.youtube.com/watch?v=zehIvTMf2mk

Latvia

1. https://www.youtube.com/watch?v=pxyQfuMZSUg
2. https://www.youtube.com/watch?v=Gpi10wyxvX4
3. https://www.youtube.com/watch?v=8Kd4INZr4e4
4. https://www.youtube.com/watch?v=USvT4lCB8ys

Netherlands

1. https://www.youtube.com/watch?v=n1-rSb3j0UM
2. https://www.youtube.com/watch?v=obCkA4YY19A
3. https://www.youtube.com/watch?v=c0YP4COuVk8
4. https://www.youtube.com/watch?v=3LqqNvpcqfA

Denmark

1. https://www.youtube.com/watch?v=31Sh_18mlI8
2. https://www.youtube.com/watch?v=YT_KefAXn7A
3. https://www.youtube.com/watch?v=3TS_JcBUHUM
4. https://www.youtube.com/watch?v=Vengv0LFYT0

Appendix 2. *Survey questions and translation*

Measured aspect	Original item (English)	Translated item (Dutch)
Shimmer	The loudness of the voice varies.	De luidheid van de stem varieert.
	The volume of the voice remains constant.	Het volume van de stem blijft constant.
Pitch variation	The height of the voice varies.	De toonhoogte van de stem varieert.
	The tone of the voice remains constant	De stemhoogte blijft constant.
Voice breaks	The speaker's voice seems to break.	De stem van de spreker lijkt te breken.
	The speaker's voice does not break.	De stem van de spreker slaat niet over.
Voice stableness (control variable)	I think this voice is stable.	Ik ervaar deze stem als stabiel.
	I do not think this is a stable voice.	Ik vind dit geen stabiele stem.

Group Process Document

Roles:

- **Introduction + Research Question**
  All searched for literature. Kim and Kirsten made it coherent and have written it.
- **Methods**
- Describe Dataset - All have contributed to data selection and thought about which countries to select. Isabeau wrote the Dataset section.
- Method 1: Acoustics - Isabeau
- Method 2: Perception - Kim & Kirsten
- Method 3: Acoustics - Ozan
- **Results**
- Result 1: Isabeau
- Result 2: Kim & Kirsten
- Result 3: Ozan
- **Discussion:** Individual
- **Reference List:** All
- **Revision**: Isabeau & Kirsten
- **Formatting**: Kim

Planning:

4/15: Finding topic and literature
4/17: First group meeting on Canvas.
Follow-up meeting: distribution of tasks and methods.
4/24: Deadline collection of videos
4/28: Group meeting at 4 PM.
We decide if we can find enough videos that are corona related (yes) and decide on what components of voice stableness we will analyze (Shimmer, Pitch and Voice breaks)
4/29: Second Group meeting on Canvas.
Teodora dropped the course, so we need to renew the roles.
5/1: Start doing our own analysis
5/8: Lecture Annotation, coding & analysis
5/9: Start writing our method and result sections
5/13: Third group meeting on Canvas.
Ozan needs to change his Method from expert coding to acoustics.
5/18: Deadline for group part of the assignment